JPL Informatics Team Monthly Report

April 2009

http://cancer.jpl.nasa.gov (informatics center)

http://edrn.nci.nih.gov/ (ops)

http://nci-test-ws04-2-v.nci.nih.gov/(test)

http://edrn.jpl.nasa.gov/ (beta)

JPL EDRN Informatics Fever Chart

EDRN Informatics Tasks	Schedule				Technical				Task Status
	Jan	Feb	Mar	Apr	Jan	Feb	Mar	Apr	
Overall	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	
Technology and Software Development									
Development									
EKE Development	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	Migrating to Plone 3.
EKE Releases and Upgrades	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	Deployed to operations at NCI
Biomarker Database Development	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	Improving audit capability, working on defining role-based groups, correcting minor bugs as found
Biomarker Database Releases and Upgrades	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	Ongoing bi-weekly releases to support curation efforts
eCAS Development	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	eCAS Curator web interface. PubMed Import similar to BMDB
eCAS Releases and Upgrades	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	Continue to ingest and curate WHI Data - Bill Bigbee. Align eCAS metadata with joint DMCC-JPL developed metadata
ERNE Releases and Upgrades	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	Identifying data quality issues.
Standards									
Ontology Development and CDEs	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	Coordinating data model and integration with CDE Development
Operations Plan/Process Development	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	Developing and refining processes and standards for capturing biomarker data across EDRN (Biomarkers, Studies, Publications, Research Data, etc).
Data Engineering & Operations									
Biomarker Data Management									
Content Management for Biomarker DB	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	Populating Biomarker Database. Communicating with DMCC, NCI and EDRN PIs. Drafting Review Process.
Content Management for eCAS	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	Updating metadata based on eCAS calls w/DMCC and NCI. Communicating with DMCC, NCI and several PIs.
eSIS Content Management WG	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	Working on RDF nightly update.
System Admin. & System Maintenance	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	Installed Tomcat WebDAV server on tumor/cancer for data delivery. Purchased 2 TB data brick (mobile eCAS) to facilitate data transfers for sites with large datasets (e.g. Dr. Liu's
Interactions									
Interactions w/outside groups	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	Interactions with several EDRN Investigators related to eCAS, BMDB and ERNE. Providing access to SELDI data for Jun Luo at Systems Analytic Inc.
Publications/Presentations	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	
Informatics Planning WG	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	Conference calls bi-monthly to discuss projects and informatics planning.

Note: Double click on the above table to open the chart in Excel.

Cautionary Items

<u>DETAILED DESCRIPTION</u>: (for items identified as yellow or red)

NONE

EKE Portal

- Accomplishments and Status
  - Resolved security issues for new public portal
  - Deployed new EDRN Public Portal to operations with Knowledge Environment
  - Initiated migration to Plone 3
  - Deployed new RDF data source: registered people

- Plans and Next Steps
  - Implement fine-grained security for biomarkers
  - Complete migration to Plone 3

Biomarker Database

- Accomplishments and Status
  - Deployed BMDB 0.8.2 release (03/27)
    - Improved integration with the Public Portal via RDF export mechanism
  - Deployed BMDB 0.8.3 release (04/30)
    - Improved sensitivity/specificity entry by supporting the post-creation modification of existing entries
- Plans and Next Steps
  - Increase the scope of the audit capabilities to capture all changes made to Biomarkers and related information in the database.
  - Create user-specific LDAP accounts and associate users with role-based groups
  - Improve support for biomarker Panels by associating panels with a Biomarker-Organ pair instead of only with a Biomarker

eCAS

- Accomplishments and Status
  - eCAS operational at http://cancer.jpl.nasa.gov/ecas
  - Deployed eCAS release 0.4.2 (3/27)
    - Inclusion of UPCI WHI data from Bigbee
  - Deployed eCAS release 0.4.1 (3/06)
    - Single Sign On capability with portal
    - Curated Hanash ROC curve data
  - Deployed eCAS release 0.4.0 (2/19)
    - Initial WHI data loaded from Tainsky's lab
    - Data Disclaimer added
    - Pagination

- Plans and Next Steps
  - Deployment of eCAS release 0.5.0
    - PubMed data import similar to BMDB
    - eCAS curator web interface
    - Curate rest of WHI data and data received from Alvin Liu
    - Align eCAS dataset metadata with joint DMCC–JPL developed metadata

ERNE

- Accomplishments and Status
    - Deployed ERNE query handler to UCSD
    - Monitored and diagnosed data quality issues

- Plans and Next Steps
    - Deploy CDC query handler
    - Automate QA

Biomarker Data Management (collaboration w/ Dartmouth)

- Accomplishments and Status
  - In collaboration with NCI (Christos Patriotis), worked with Dartmouth on Biomarker Database content management
  - Continue to curate biomarkers providing deep content such as publications, alternate names, etc. (Ovary, Liver, Lung, Bladder)
    - Populating Sensitivity and Specificity for each marker
  - Contacted several EDRN Investigators for Biomarker content (Dr. Morrero, Dr. Brenner, DMCC and NCI)
  - Conducted the following conference calls:
    - Bi-weekly calls to discuss the Biomarker Database curation, system development and process development and evaluation of biomarkers with NCI and the DMCC
    - Bi-weekly calls to discuss eSIS curation, process development and data modeling for eSIS
  - Continue to work with the DMCC and Investigators to load WHI data, Dr. Grizzle test data, Dr. Liu data
  - Complete Protocol Records for Biomarker and eCAS content as it is curated
  - Create documentation of curation process, to ensure consistent curation between biomarkers and as the curation work expands

Biomarker Data Management Continued

- Plans and Next Steps
  - Continue curating next set of biomarkers for GI
  - Continue researching and contacting investigators to provide deep content for existing markers (populating sensitivity and specificity, etc.)
  - Continue working with Dartmouth and DMCC to provide specifications on Biomarker Database and Biomarker Data Model
  - Continue updating documentation of biomarker database curation process
  - Continue working to curate metadata for WHI Data, Dr. Grizzle test data and Dr. Liu data to eCAS
  - Continue working with EDRN sites to gather feedback on applicable metadata for eCAS
  - Continue contacting PIs to complete Protocol information
  - Continue all BMDB, eCAS, eSIS and informatics calls.

Publications, Presentations and Outreach

- Publications
  - Andrew.F.Hart, Chris A. Mattmann, John J. Tran, Daniel J. Crichton, J. Steven Hughes, Heather Kincaid, Sean Kelly, Kristen Anton, Donald Johnsey, Christos Patriotis. Enabling Effective Curation of Cancer Biomarker Research Data. Submitted to CBMS 2009.

- Meetings
  - Conference call with Dr Hannes Niedner to present and demonstrate his grid service for ERNE based on caBIG.

EDRN Meetings/Interactions

- Informatics Planning Meetings (bi-weekly)
  - Thursdays, 11 AM PT
- ERNE teleconferences (Monthly)
  - Mondays, 10 AM PT
- Biomarker Curation Teleconference (bi-weekly)
  - Thursdays, 8:30 AM PT
- eSIS Coordination Teleconference (bi-weekly)
  - Tuesdays, 8:30 AM PT
- eCAS Coordination Teleconference (bi-weekly)
  - Tuesdays, 8:30 AM PT
- JPL Informatics Teleconference (weekly)
  - Tuesdays, 10 AM PT
- NYU Informatics Teleconference (weekly)
  - Tuesdays, 8 AM PT

NCI/JPL Items

- Portal: transition to operations
  - If test server checks out, need to coordinate switchover to http://edrn.nci.nih.gov/ (this is expected in mid January)
- Need to coordinate and continue to establish informatics data sharing policies
  - Recommend some discussion between NCI, JPL and DMCC for specific language on data sharing for the EDRN renewal
- JPL is interested in coordinating three face-to-face meetings in 2009
  - Technical Meeting on Informatics (Agenda Drafted – tentative Dates Jan. 29 and 30)
  - Biomarker Data Management Review (~March 2009)
  - EDRN Data Providers and Users Meeting (~October 2009)

Links, etc

- EDRN operational website
  http://edrn.nci.nih.gov/

- EDRN beta website
  http://edrn.jpl.nasa.gov/

- NCI test website
  http://nci-test-ws04-2-v.nci.nih.gov/

- JPL EDRN Informatics Center Website
  http://cancer.jpl.nasa.gov/

- eCAS
  http://cancer.jpl.nasa.gov/ecas

- BMDB curation system
  http://cancer.jpl.nasa.gov/bmdb

Informatics Personnel

- Dan Crichton, PI
- Chris Mattmann, Development Lead
- Heather Kincaid, Operations Lead
- Sean Kelly, Technologist
- Steve Hughes, Information Architect
- Andrew Hart, Developer
- John Tran, Developer
- Thuy Tran, Developer